

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



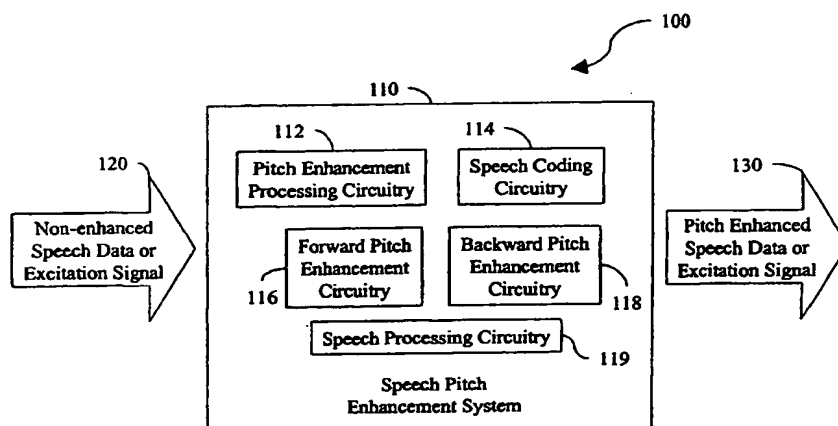
(43) International Publication Date  
11 January 2001 (11.01.2001)

PCT

(10) International Publication Number  
**WO 01/03125 A1**

- (51) International Patent Classification?: **G10L 21/02**, 19/12 (72) Inventor: **GAO, Yang**; 26586 San Torini Road, Mission Viejo, CA 92692 (US).
- (21) International Application Number: **PCT/US00/18232** (74) Agent: **SHORT, Shayne, X.**; Akin, Gump, Strauss, Hauer & Feld, LLP, Suite 1900, 816 Congress Avenue, Austin, TX 78701 (US).
- (22) International Filing Date: **30 June 2000 (30.06.2000)**
- (25) Filing Language: **English** (81) Designated States (*national*): **CN, JP, KR.**
- (26) Publication Language: **English** (84) Designated States (*regional*): **European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).**
- (30) Priority Data:  
60/142,092 2 July 1999 (02.07.1999) US  
60/365,444 2 August 1999 (02.08.1999) US  
Published:  
— With international search report.
- (71) Applicant: **CONEXANT SYSTEMS, INC.** [US/US]; 4311 Jamboree Road, Newport Beach, CA 92660-3095 (US).  
*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: **BI-DIRECTIONAL PITCH ENHANCEMENT IN SPEECH CODING SYSTEMS**



(57) Abstract: A bi-directional pitch enhancement system for speech coding systems. As speech data applications continue to operate in areas having intrinsic bandwidth limitations, the perceptual quality of reproduced speech data in typical speech coding systems suffers significantly. The present invention employs forward pitch enhancement and backward pitch enhancement to maintain a high perceptual quality in reproduced speech. If desired, the backward pitch enhancement is generated using the forward pitch enhancement itself with the backward pitch enhancement being a mirror image of the forward pitch enhancement that was previously generated. Alternatively, in other embodiments of the invention, the backward pitch enhancement is generated independent of the forward pitch enhancement. The backward pitch enhancement is usually performed on the fixed codebook in code excited linear prediction (CELP) or is performed as post-processing in the decoder.

WO 01/03125 A1

**TITLE:       BI-DIRECTIONAL PITCH ENHANCEMENT  
              IN SPEECH CODING SYSTEMS**

**SPECIFICATION**

**CROSS-REFERENCE TO RELATED APPLICATIONS**

The present application claims priority to U.S. Provisional Application Serial No. 60/142,092, entitled "Bi-Directional Pitch Enhancement In Speech Coding Systems," (Attorney Docket No. 97RSS380P), filed July 2, 1999 and U.S. Application Serial No. 09/365,444, entitled "Bi-Directional Pitch Enhancement In Speech Coding Systems," (Attorney Docket No. 97RSS380), filed August 2, 1999.

**BACKGROUND**

**1.     Technical Field**

The present invention relates generally to speech coding; and, more particularly, it relates to low bit rate speech coding systems that employ pitch enhancement to improve the perceptual quality of reproduced speech.

**2.     Description of Related Art**

Conventional speech coding systems typically employ only forward pitch enhancement in code-excited linear prediction speech coding systems. This is largely due to the fact that the sub-frame size of conventional speech codecs, having relatively large bandwidth availability, can provide sufficient perceptual quality with forward pitch enhancement alone. However, for lower bit rates within various communication media

employed in speech coding systems, the perceptual quality of reproduced speech, after synthesis, fails to maintain a high perceptual quality.

For conventional speech coding systems that operate at these decreased bit rates, the pitch lag, that is generated during pitch prediction, is commonly much shorter than the overall sub-frame size, i.e., it covers a relatively small portion of the overall sub-frame. This characteristic is more accentuated for those speakers having a higher (shorter) pitch, such as females and children. Traditional excitation codebook structures do not afford a sufficient high perceptual quality when operating at low bit rates. This is primarily because the periodicity of the voiced signal is not sufficiently established, or the excitation vector extracted from the codebook is insufficiently rich to generate a synthesized speech signal having a high perceptual quality.

As the sub-frame size of speech coding systems becomes larger, as is commonly associated with communication systems that have decreasing bit rates, the fact that pitch enhancement is performed in only the forward direction results in significantly poorer perceptual quality. This is due, among other reasons, to the fact that there is a significant amount of dead space in the sub-frame due to the absence of many pulses. In conventional speech coding systems that operate at higher bit rate, having consequently shorter sub-frames, this effect is not typically audibly perceived by the human ear. This effect of lower perceptual quality is realized in nearly all speech coding systems that deal with speech coding having relatively low available bit rates.

Further limitations and disadvantages of conventional and traditional systems will become apparent to one of skill in the art through comparison of such systems with the present invention as set forth in the remainder of the present application with reference to the drawings.

### **SUMMARY OF THE INVENTION**

Various aspects of the present invention can be found in a speech coding system that employs forward pitch enhancement and backward pitch enhancement. In certain embodiments of the invention, the forward pitch enhancement and the backward pitch enhancement are performed in a single portion of the entire speech coding system. For example, in speech coding systems having a speech codec, wherein the speech codec contains an encoder and a decoder, the forward pitch enhancement and the backward pitch enhancement are performed in both the encoder and the decoder of the speech codec. Alternatively, in other embodiments of the invention, the forward pitch enhancement and the backward pitch enhancement are performed only in the decoder of the speech codec. As determined by the specific application, the forward pitch enhancement and the backward pitch enhancement are performed in a distributed manner, each being performed, at least in part, in each one of the encoder and the decoder of the speech codec.

In certain embodiments of the invention, the backward pitch enhancement is generated using the forward pitch enhancement itself. The backward pitch enhancement is a mirror image of the forward pitch enhancement that is previously generated; the backward pitch enhancement is generated dependent on the forward pitch enhancement. Alternatively, in other embodiments of the invention, the backward pitch enhancement is generated independent of the forward pitch enhancement; the backward pitch enhancement is generated irrespective of the forward pitch enhancement that has previously been generated.

The speech coding system, built in accordance with the present invention, is appropriately geared toward those speech coding systems that operate using communication media having limited or constrained bandwidth availability. Any communication media may be employed within in the invention, without departing from the scope and spirit thereof. Examples of such communication media include, but are not limited to, wireless

communication media, wire-based telephonic communication media, fiber-optic communication media, and ethernet.

Other aspects, advantages and novel features of the present invention will become apparent from the following detailed description of the invention when considered in conjunction with the accompanying drawings.

### **BRIEF DESCRIPTION OF DRAWINGS**

Fig. 1 is a system diagram illustrating one embodiment of a speech pitch enhancement system built in accordance with the present invention.

Fig. 2 is a system diagram illustrating one embodiment of a distributed speech codec that employs speech pitch enhancement in accordance with the present invention.

Fig. 3 is a system diagram illustrating another embodiment of a distributed speech codec that employs speech pitch enhancement in accordance with the present invention.

Fig. 4 is a system diagram illustrating another embodiment of an integrated speech codec that employs speech pitch enhancement in accordance with the present invention.

Fig. 5 is a diagram illustrating a speech sub-frame depicting forward and backward predicted pulses to perform pitch enhancement in accordance with the present invention.

Fig. 6 illustrates a functional block diagram illustrating an embodiment of the present invention that generates backward speech pitch enhancement using forward speech pitch enhancement in accordance with the present invention.

Fig. 7 illustrates a functional block diagram illustrating an embodiment of the present invention that performs backward speech pitch enhancement independent of forward speech pitch enhancement in accordance with the present invention.

### **DETAILED DESCRIPTION OF DRAWINGS**

Fig. 1 is a system diagram illustrating one embodiment 100 of a speech pitch enhancement system 110 built in accordance with the present invention. The speech pitch enhancement system 110 contains, among other things, pitch enhancement processing circuitry 112, speech coding circuitry 114, forward pitch enhancement circuitry 116, backward pitch enhancement circuitry 118, and speech processing circuitry 119. The speech pitch enhancement system 110 operates on non-enhanced speech data or excitation signal 120 and generates pitch enhanced speech data 130. The pitch enhanced speech data or excitation signal 130 contains speech data having pitch prediction and pitch enhancement performed in both the forward and backward directions with respect to a speech sub-frame. The speech pitch enhancement system 110 operates only on an excitation signal in certain embodiments of the invention, and the speech pitch enhancement system 110 operates only on speech data in other embodiments of the invention.

In certain embodiments of the invention, the speech pitch enhancement system 110 operates independently to generate backward pitch prediction using the backward pitch enhancement circuitry 118. Alternatively, the forward pitch enhancement circuitry 116 and the backward pitch enhancement circuitry 118 operate cooperatively to generate the overall pitch enhancement of the speech coding system. A supervisory control operation, monitoring the forward pitch enhancement circuitry 116 and the backward pitch enhancement circuitry 118, is performed using the pitch enhancement processing circuitry 112 in other embodiments of the invention. The speech processing circuitry 119 includes, but is not limited to, that speech processing circuitry known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. The speech coding circuitry 114 similarly includes, but is not limited to, circuitry known to those of skill in the art of speech coding. Such speech coding known to those having skill in the art includes,

among other speech coding methods, code-excited linear prediction, algebraic code-excited linear prediction, and pulse-like excitation.

Fig. 2 is a system diagram illustrating one embodiment of a distributed speech codec 200 that employs speech pitch enhancement in accordance with the present invention. A speech encoder 220 of the distributed speech codec 200 performs pitch enhancement coding 221. The pitch enhancement coding 221 is performed using both backward pulse pitch prediction circuitry 222 and forward pulse pitch prediction circuitry 223. As described above in another embodiment of the invention, the pitch enhancement coding 221 generates pitch prediction and pitch enhancement in both the forward and backward directions within the speech sub-frame. The speech encoder 220 of the distributed speech codec 200 also performs main pulse coding 225 of a speech signal including both sign coding 226 and location coding 227 within a speech sub-frame. Speech processing circuitry 229 is also employed within the speech encoder 220 of the distributed speech codec 200 to assist in speech processing using methods known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. Additionally, the speech processing circuitry 229 operates cooperatively with the backward pulse pitch prediction circuitry 222 and forward pulse pitch prediction circuitry 223 in certain embodiments of the invention. The speech data, after having been processed, at least to some extent by the speech encoder 220 of the distributed speech codec 200 is transmitted via a communication link 210 to a speech decoder 230 of the distributed speech codec 200. The communication link 210 is any communication media capable of transmitting voiced data, including but not limited to, wireless communication media, wire-based telephonic communication media, fiber-optic communication media, and ethernet. Any communication media capable of transmitting speech data is included in the communication link 210 without departing from the scope and spirit of the invention. The speech decoder 230 of the distributed speech codec 200 contains,



among other things, speech reproduction circuitry 232, perceptual compensation circuitry 234, and speech processing circuitry 236.

In certain embodiments of the invention, the speech processing circuitry 229 and the speech processing circuitry 236 operate cooperatively on the speech data within the entirety of the distributed speech codec 200. Alternatively, the speech processing circuitry 229 and the speech processing circuitry 236 operate independently on the speech data, each serving individual speech processing functions in the speech encoder 220 and the speech decoder 230, respectively. The speech processing circuitry 229 and the speech processing circuitry 236 include, but are not limited to, that speech processing circuitry known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. The main pulse coding circuitry 225 similarly includes, but is not limited to, circuitry known to those of skill in the art of speech coding. Examples of such main pulse coding circuitry 225 include that circuitry known to those having skill in the art, among other main pulse coding methods, code-excited linear prediction, algebraic code-excited linear prediction, and pulse-like excitation, as described above in another embodiment of the invention.

Fig. 3 is a system diagram illustrating another embodiment of a distributed speech codec 300 that employs speech pitch enhancement in accordance with the present invention. A speech encoder 320 of the distributed speech codec 300 performs main pulse coding 325 of a speech signal including both sign coding 326 and location coding 327 within a speech sub-frame. Speech processing circuitry 329 is also employed within the speech encoder 320 of the distributed speech codec 300 to assist in speech processing using methods known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. The speech data, after having been processed, at least to some extent by the speech encoder 320 of the distributed speech codec 300 is transmitted via a communication link 310 to a speech decoder 330 of the distributed speech codec 300. The communication link 310 is

any communication media capable of transmitted voiced data, including but not limited to, wireless communication media, wire-based telephonic communication media, fiber-optic communication media, and ethernet. Any communication media capable of transmitting speech data is included in the communication link 310 without departing from the scope and spirit of the invention. A speech decoder 330 of the distributed speech codec 300 performs pitch enhancement coding 321. The pitch enhancement coding 321 is performed using both backward pulse pitch prediction circuitry 322 and forward pulse pitch prediction circuitry 323. As described above in various embodiments of the invention, the pitch enhancement coding 321 generates pitch prediction and pitch enhancement in both the forward and backward directions within the speech sub-frame. Speech processing circuitry 336 is also employed within the speech decoder 330 of the distributed speech codec 300 to assist in speech processing using methods known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. Additionally, the speech processing circuitry 339 operates cooperatively with the backward pulse pitch prediction circuitry 322 and forward pulse pitch prediction circuitry 323 in certain embodiments of the invention.

In certain embodiments of the invention, the speech processing circuitry 329 and the speech processing circuitry 336 operate cooperatively on the speech data within the entirety of the distributed speech codec 300. Alternatively, the speech processing circuitry 329 and the speech processing circuitry 336 operate independently on the speech data, each serving individual speech processing functions in the speech encoder 320 and the speech decoder 330, respectively. The speech processing circuitry 329 and the speech processing circuitry 336 include, but are not limited to, that speech processing circuitry known to those having skill in the art of speech processing to operate on and perform manipulation of speech data. The main pulse coding circuitry 325 similarly includes, but is not limited to, circuitry known to those of skill in the art of speech coding. Examples of such main pulse coding circuitry

325 includes that circuitry known to those having skill in the art, among other main pulse coding methods, code-excited linear prediction, algebraic code-excited linear prediction, and pulse-like excitation, as described above in another embodiment of the invention.

Fig. 4 is a system diagram illustrating another embodiment 400 of an integrated speech codec 420 that employs speech pitch enhancement in accordance with the present invention. The integrated speech codec 420 contains, among other things, a speech encoder 422 that communicates with a speech decoder 424 via a low bit rate communication link 410. The low bit rate communication link 410 is any communication media capable of transmitting voiced data, including but not limited to, wireless communication media, wire-based telephonic communication media, fiber-optic communication media, and ethernet. Any communication media capable of transmitting speech data is included in the low bit rate communication link 410 without departing from the scope and spirit of the invention. Pitch enhancement coding 421 is performed in the integrated speech codec 420. The pitch enhancement coding 421 is performed using, among other things, backward pulse pitch prediction circuitry 422 and forward pulse pitch prediction circuitry 423. As described above in various embodiments of the invention, the backward pulse pitch prediction circuitry 422 and the forward pulse pitch prediction circuitry 423 operate cooperatively in certain embodiments of the invention, and independently in other embodiments of the invention.

As shown in the embodiment 400, the backward pulse pitch prediction circuitry 422 and the forward pulse pitch prediction circuitry 423 are contained within the entirety of the integrated speech codec 420. If desired, the backward pulse pitch prediction circuitry 422 and the forward pulse pitch prediction circuitry 423 are both contained in each of the speech encoder 422 and the speech decoder 424 in certain embodiments of the invention. Alternatively, either one of the backward pulse pitch prediction circuitry 422 or the forward pulse pitch prediction circuitry 423 is contained in only one of the speech encoder 422 and

the speech decoder 424 in other embodiments of the invention. Depending on the specific application at hand, a user can select to place the backward pulse pitch prediction circuitry 422 and the forward pulse pitch prediction circuitry 423 in only one or either of the speech encoder 422 and the speech decoder 424. Various embodiments are envisioned in the invention, without departing from the scope and spirit thereof, to place various amounts of the backward pulse pitch prediction circuitry 422 and the forward pulse pitch prediction circuitry 423 in the speech encoder 422 and the speech decoder 424. For example, a predetermined portion of the backward pulse pitch prediction circuitry 422 is placed in the speech encoder 422 while a remaining portion of the backward pulse pitch prediction circuitry 422 is placed in the speech decoder 424 in certain embodiments of the invention. Similarly, a predetermined portion of the forward pulse pitch prediction circuitry 423 is placed in the speech encoder 422 while a remaining portion of the forward pulse pitch prediction circuitry 423 is placed in the speech decoder 424 in certain embodiments of the invention.

Fig. 5 is a coding diagram 500 illustrating a speech sub-frame 510 depicting forward pitch enhancement and backward pitch enhancement performed in accordance with the present invention. A main pulse  $M_0$  520 is generated in the speech sub-frame 510 using any method known to those having skill in the art of speech processing, including but not limited to, code-excited linear prediction, algebraic code-excited linear prediction, analysis by synthesis speech coding, and pulse-like excitation. Using various methods of speech processing, including those methods described above that are employed in various embodiments of the invention, a forward predicted pulse  $M_1$  530, a forward predicted pulse  $M_2$  540, and a forward predicted pulse  $M_3$  550 are all generated and placed within the speech sub-frame 510. As described above, the generation of the forward predicted pulse  $M_1$  530, the forward predicted pulse  $M_2$  540, and the forward predicted pulse  $M_3$  550 is performed

using various processing circuitry in certain embodiments of the invention. In addition, a backward predicted pulse  $M_{-1}$  560 and a backward predicted pulse  $M_{-2}$  570 are also generated in accordance with the invention.

In certain embodiments of the invention, the backward predicted pulse  $M_{-1}$  560 and the backward predicted pulse  $M_{-2}$  570 are generated using the forward predicted pulse  $M_1$  530, the forward predicted pulse  $M_2$  540, and the forward predicted pulse  $M_3$  550. Alternatively, in other embodiments of the invention, the backward predicted pulse  $M_{-1}$  560 and the backward predicted pulse  $M_{-2}$  570 are generated independent of the forward predicted pulse  $M_1$  530, the forward predicted pulse  $M_2$  540, and the forward predicted pulse  $M_3$  550. An example of independent generation of the backward predicted pulse  $M_{-1}$  560 and the backward predicted pulse  $M_{-2}$  570 is an implementation within software wherein the time scale of the speech sub-frame 510 is reversed in software. The main pulse  $M_0$  520 is used in a similar manner to generate both the forward predicted pulse  $M_1$  530, the forward predicted pulse  $M_2$  540, and the forward predicted pulse  $M_3$  550, and the backward predicted pulse  $M_{-1}$  560 and the backward predicted pulse  $M_{-2}$  570. That is to say, the process is performed once in the typical forward direction, and after the speech sub-frame 510 is reversed in software, the process is performed once again in the atypical backward direction, yet it employs the same mathematical method, i.e., only the data are reversed with respect to speech sub-frame 510.

Fig. 6 illustrates a functional block diagram illustrating an embodiment 600 of the present invention that generates backward speech pitch enhancement using forward speech pitch enhancement in accordance with the present invention. In a block 610, a speech signal is processed. In a block 620, a main pulse of the speech data is coded. In an alternative process block 655, the speech data information is transmitted via a communication link. The alternative process block 655 is employed in embodiments of the invention wherein the

forward pitch enhancement and backward pitch enhancement are performed after the coded speech data is transmitted for speech reproduction. In a block 630, forward pitch enhancement is performed, and in a block 640, backward pitch enhancement is performed. The backward pitch enhancement of the block 640 is a mirror image of the forward pitch enhancement that is generated in the block 630 in certain embodiments of the invention. In other embodiments, the backward pitch enhancement of the block 640 is not a mirror image of the forward pitch enhancement that is generated in the block 630. In an alternative process block 650, the speech data information is transmitted via a communication link. The alternative process block 650 is employed in embodiments of the invention wherein the forward pitch enhancement and backward pitch enhancement are performed prior to the coded speech data being transmitted for speech reproduction. In a block 660, the speech signal is reconstructed/synthesized.

In certain embodiments of the invention, the backward pitch enhancement performed in the block 640 is simply a duplicate of the forward pitch enhancement performed in the block 630, i.e., backward pitch enhancement of the block 640 is a mirror image of the forward pitch enhancement generated in the block 630. For example, after the forward pitch enhancement is performed in the block 630, the resultant pitch enhancement is simply copied and reversed within a speech sub-frame to generate the backward pitch enhancement performed in the block 640 using any method known to those skilled in the art of speech processing for synthesizing and reproducing a speech signal.

Fig. 7 illustrates a functional block diagram illustrating an embodiment 700 of the present invention that performs backward speech pitch enhancement independent of forward speech pitch enhancement in accordance with the present invention. In a block 710, a speech signal is processed. In a block 720, a main pulse of the speech data is coded. In an alternative process block 755, the speech data information is transmitted via a communication

link. The alternative process block 655 is employed in embodiments of the invention wherein the forward pitch enhancement and backward pitch enhancement are performed after the coded speech data is transmitted for speech reproduction. In a block 730, forward pitch enhancement is performed, and in a block 740, backward pitch enhancement is performed. The backward pitch enhancement of the block 740 is performed after the speech data is reversed; the backward pitch enhancement of the block 740 is performed independently of the forward pitch enhancement that is performed that is performed in the block 730. This particular embodiment differs from that illustrated in the embodiment 600, in that, the speech data are reversed and the backward pitch enhancement of the block 740 is generated as if an entirely new set of speech data were being processed. Conversely, in the embodiment 600, the resulting pitch enhancement itself is utilized, but it extended in the reverse direction. In certain embodiments of the embodiment 700, it is as if two sets of speech data are being processed for each sub-frame; one set of data is processed to generate the pitch prediction in the forward direction in the block 730, and one set of data is processed to generate the pitch prediction in the backward direction in the block 740, yet they are both operating on the same sub-frame of speech data. In an alternative process block 750, the speech data information is transmitted via a communication link. The alternative process block 650 is employed in embodiments of the invention wherein the forward pitch enhancement of the block 730 and backward pitch enhancement of the block 740 are performed prior to the coded speech data being transmitted for speech reproduction. In a block 760, the speech signal is reconstructed/synthesized.

In view of the above detailed description of the present invention and associated drawings, other modifications and variations will now become apparent to those skilled in the art. It should also be apparent that such other modifications and variations may be effected without departing from the spirit and scope of the present invention.

**CLAIMS**

What is claimed is:

1. A speech codec that performs pitch enhancement on speech data, the speech codec comprising:

an encoder;

a communication link communicatively coupled to the encoder;

a decoder communicatively coupled to the encoder via the communication link;

a forward pitch enhancement circuit contained within the speech codec, the forward pitch enhancement circuit improves the perceptual quality of the speech data for reproduction; and

a backward pitch enhancement circuit contained within the speech codec, the backward pitch enhancement circuit improves the perceptual quality of the speech data for reproduction.

2. The speech codec of claim 1, wherein the backward pitch enhancement circuit is contained in the decoder of the speech codec.

3. The speech codec of claim 1, wherein the backward pitch enhancement circuit is distributed in the encoder and the decoder of the speech codec.

4. The speech codec of claim 1, wherein the forward pitch enhancement circuit and the backward pitch enhancement circuit operate cooperatively to improve the perceptual quality of the speech data for reproduction.



5. The speech codec of claim 1, wherein the forward pitch enhancement circuit and the backward pitch enhancement circuit operate independently to improve the perceptual quality of the speech data for reproduction.

6. The speech codec of claim 1, wherein the backward pitch enhancement circuit operates on a speech sub-frame; and

the backward pitch enhancement circuit places at least one backward predicted pulse within the speech sub-frame

7. A speech pitch enhancement system that operates on speech data, the speech pitch enhancement system comprising:

a backward pitch enhancement circuit that improves the perceptual quality of the speech data for reproduction; and

a speech processing circuit communicatively coupled to the backward pitch enhancement circuit, the speech processing circuit performs manipulation of the speech data.

8. The speech pitch enhancement system of claim 7, further comprising a forward pitch enhancement circuit communicatively coupled to the backward pitch enhancement circuit, the forward pitch enhancement circuit improves the perceptual quality of the speech data for reproduction.

9. The speech pitch enhancement system of claim 8, wherein the forward pitch enhancement circuit and the backward pitch enhancement circuit operate cooperatively to improve the perceptual quality of the speech data for reproduction.

10. The speech pitch enhancement system of claim 8, wherein the forward pitch enhancement circuit and the backward pitch enhancement circuit operate independently to improve the perceptual quality of the speech data for reproduction.

11. The speech pitch enhancement system of claim 7, further comprising a speech codec, the speech codec comprising an encoder and a decoder; and

wherein the backward pitch enhancement circuit is contained within at least one of the encoder and the decoder.

12. The speech pitch enhancement system of claim 7, further comprising a speech codec, the speech codec comprising an encoder and a decoder; and

wherein the backward pitch enhancement circuit is distributed between the encoder and the decoder.

13. The speech pitch enhancement system of claim 7, further comprising a speech codec, the speech codec comprises an encoder, a decoder, and a communication link, the communication link communicatively coupled the encoder and the decoder;

wherein the backward pitch enhancement circuit is distributed in the encoder and the decoder;

the speech data comprises a first portion of the speech data and a second portion of the speech data;

the perceptual quality of the first portion of the speech data is improved in the encoder using the backward pitch enhancement circuit;

the first portion of the speech data and the second portion of the speech data are transmitted from the encoder to the decoder via the communication link; and

the perceptual quality of the second portion of the speech data is improved in the decoder using the backward pitch enhancement circuit.

14. The speech pitch enhancement system of claim 7, wherein the speech pitch enhancement system employs code-excited linear prediction.

15. The speech pitch enhancement system of claim 7, wherein the backward pitch enhancement circuit operates on a speech sub-frame; and

the backward pitch enhancement circuit places at least one backward predicted pulse within the speech sub-frame.

16. A method that performs speech pitch enhancement on at least one of speech data and an excitation signal, the method comprising:

performing forward pitch enhancement on the at least one of the speech data and the excitation signal; and

performing backward pitch enhancement on the at least one of the speech data and the excitation signal.

17. The method of claim 16, wherein the performing forward pitch enhancement on the at least one of the speech data and the excitation signal and the performing backward pitch enhancement on the at least one of the speech data and the excitation signal are performed independently to improve the perceptual quality of the speech data for reproduction.

18. The method of claim 16, wherein the performing forward pitch enhancement on the at least one of the speech data and the excitation signal and the performing backward pitch

enhancement on the at least one of the speech data and the excitation signal are performed cooperatively to improve the perceptual quality of the speech data for reproduction.

19. The method of claim 16, wherein the performing forward pitch enhancement on the at least one of the speech data and the excitation signal and the performing backward pitch enhancement on the at least one of the speech data and the excitation signal are performed using a speech codec.

20. The method of claim 16, further comprising performing code-excited linear prediction.

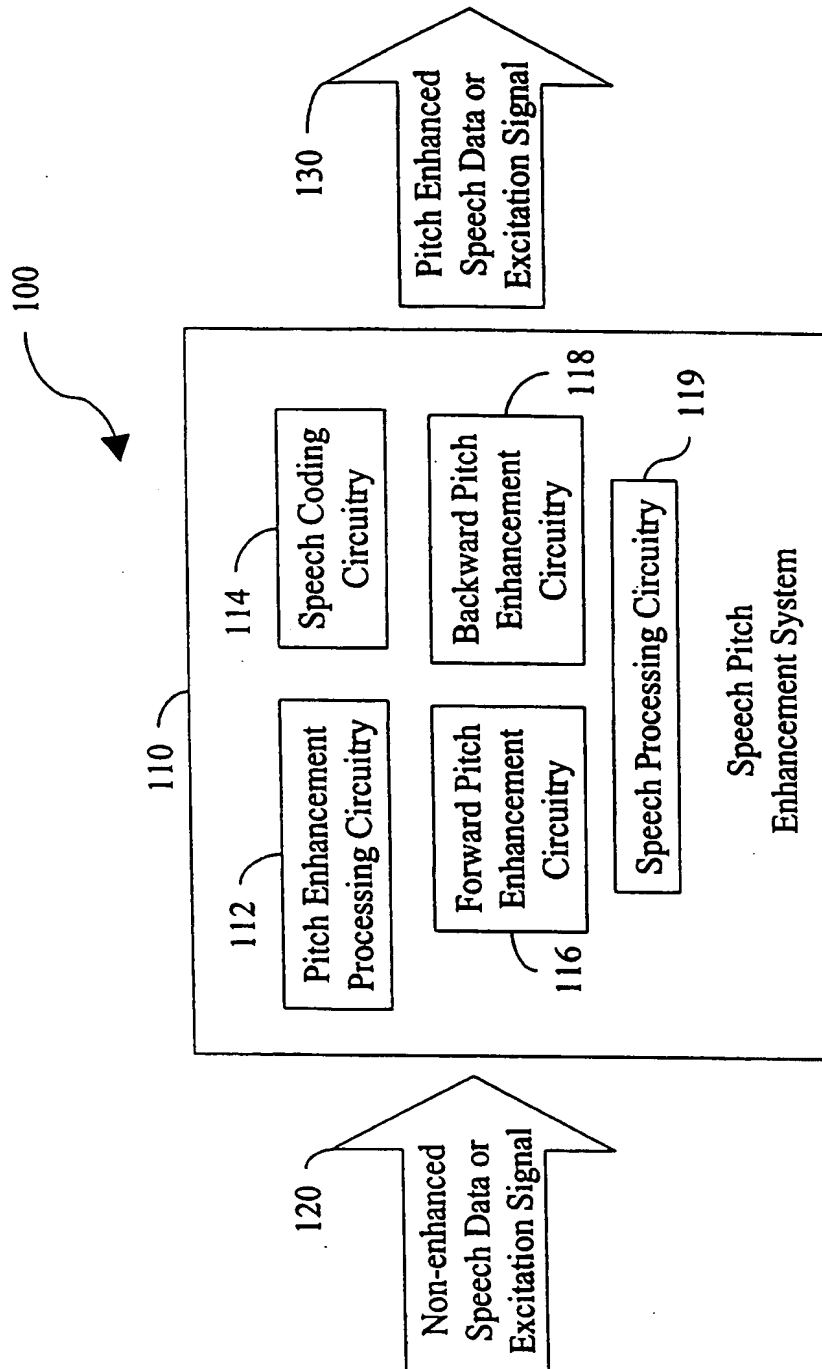


Fig. 1

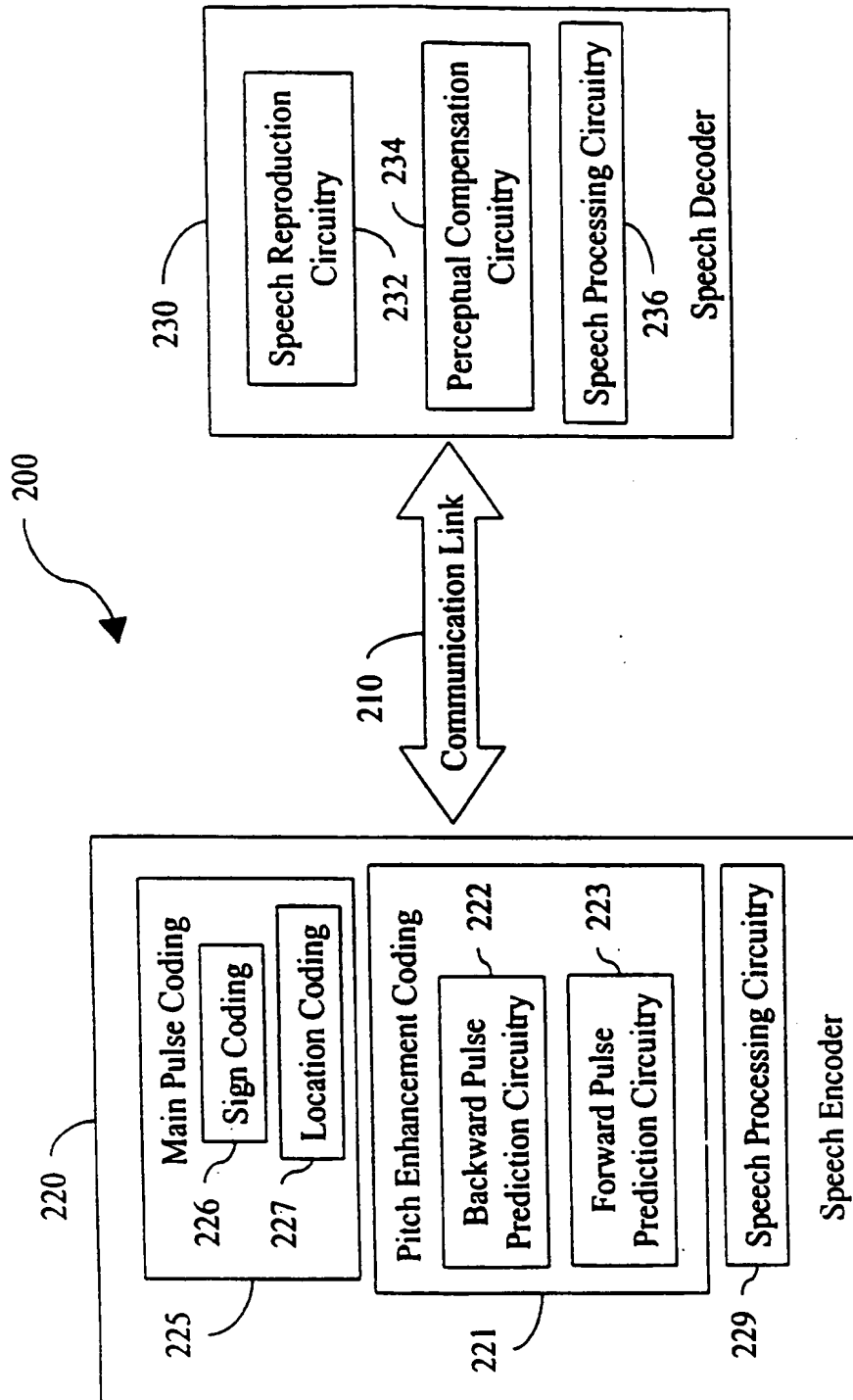


Fig. 2

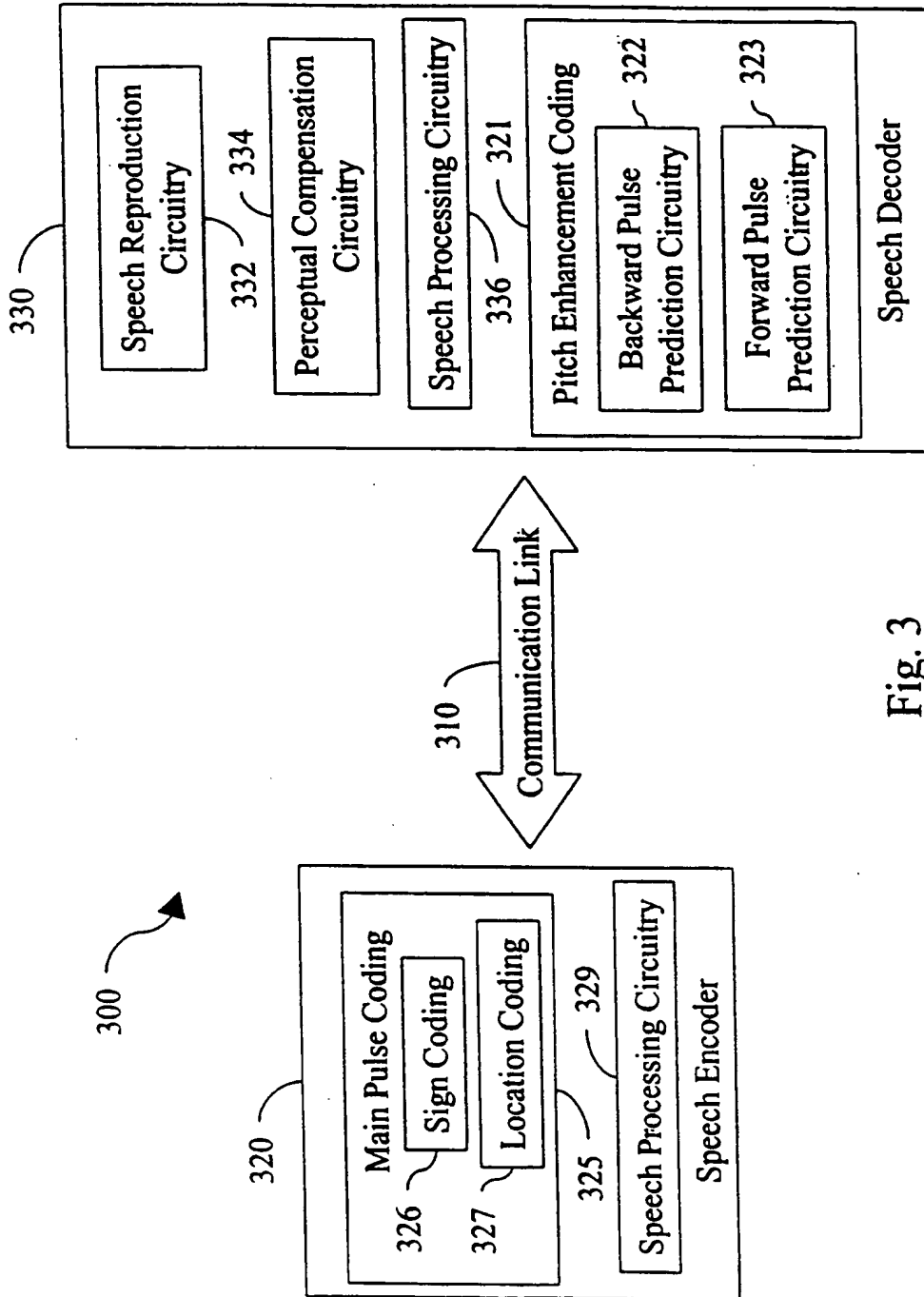


Fig. 3

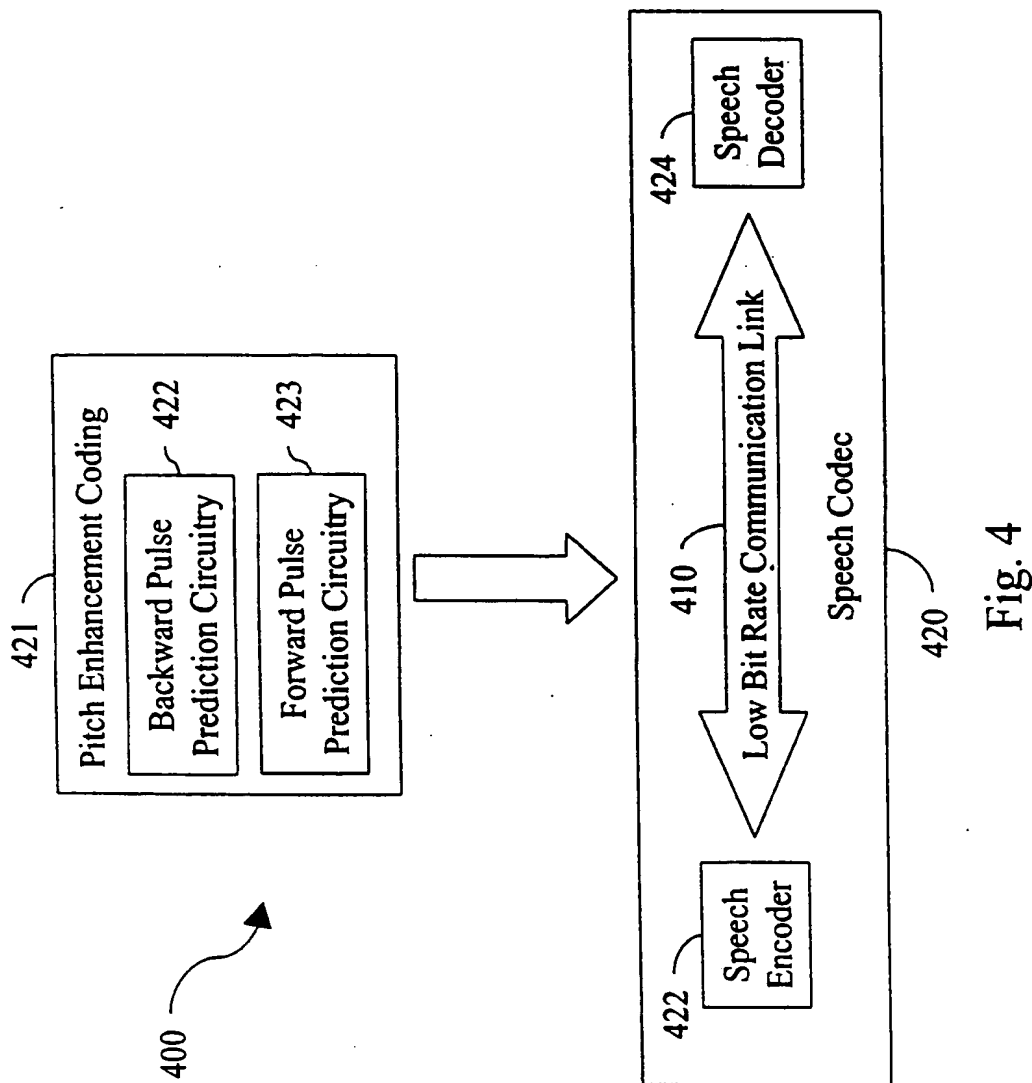


Fig. 4



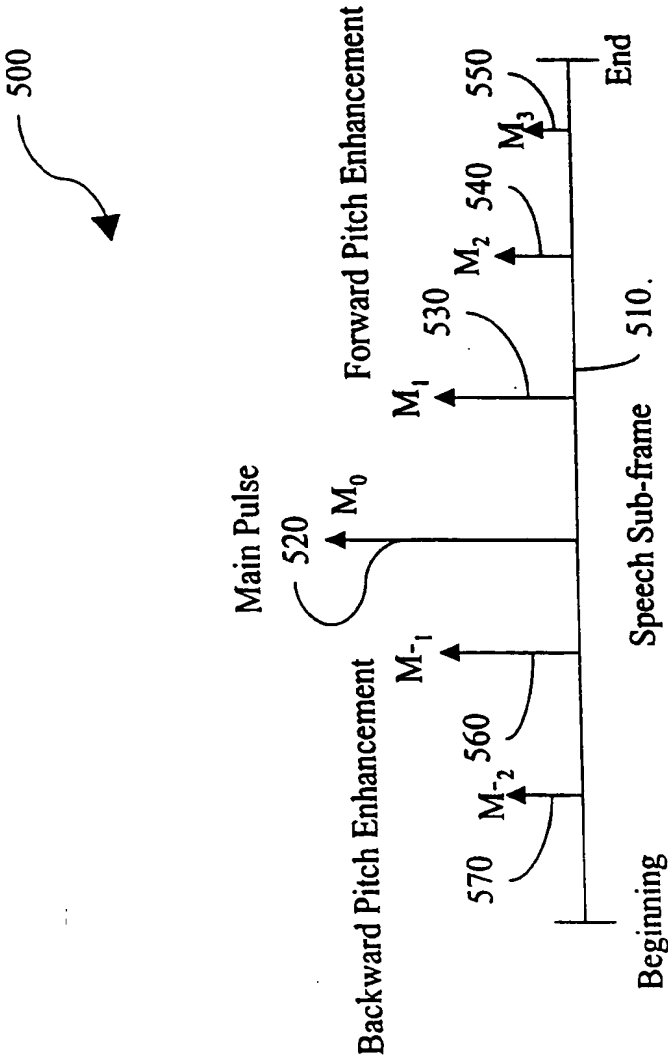


Fig. 5

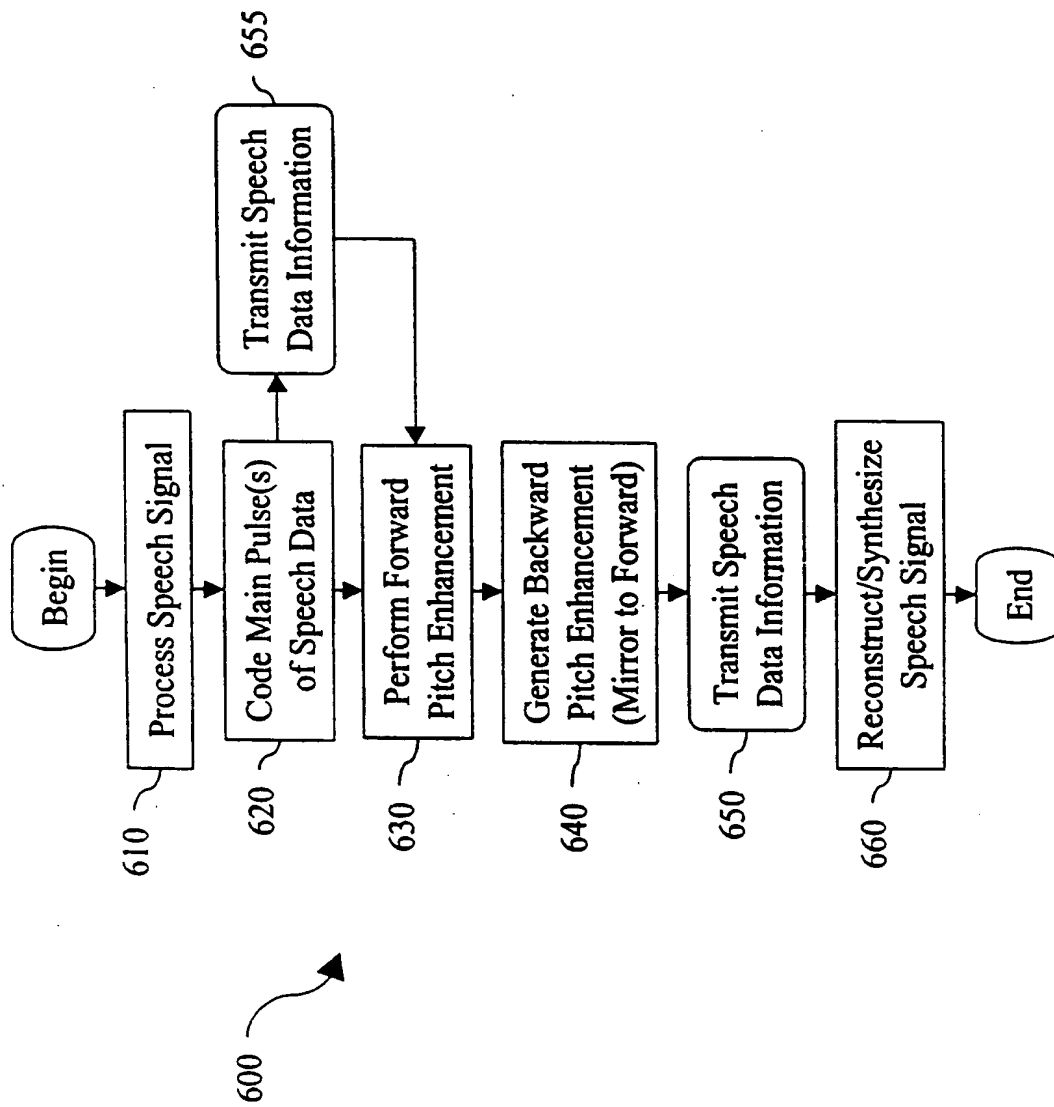


Fig. 6

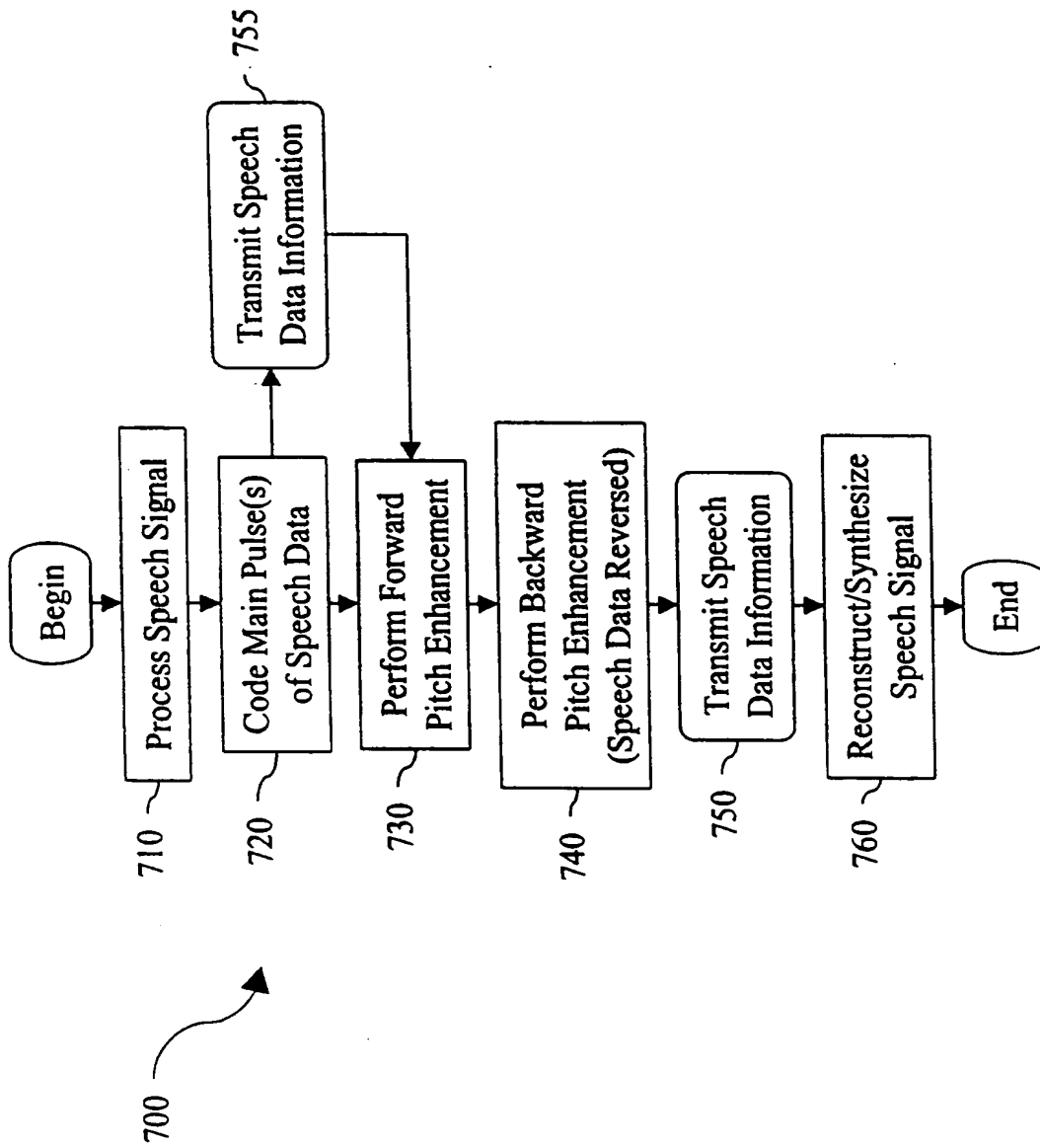


Fig. 7

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/18232

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G10L21/02 G10L19/12

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, INSPEC, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LI H ET AL: "NON-LINEAR TECHNIQUES FOR PITCH AND WAVEFORM ENHANCEMENT IN PWI CODERS" IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), 21 April 1997 (1997-04-21), pages 1563-1566, XP000822759 ISBN: 0-8186-7920-4 abstract page 1564, right-hand column page 1565	1,4,7,9, 16,18
A	WO 97 27578 A (MOTOROLA INC) 31 July 1997 (1997-07-31) page 21, line 38 -page 22, line 3 page 24, line 1 - line 17  -/-	1-20

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

4 October 2000

Date of mailing of the international search report

18/10/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Ramos Sánchez, U

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/18232

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 926 660 A (TOKYO SHIBAURA ELECTRIC CO) 30 June 1999 (1999-06-30) column 10, line 39 - line 50 column 12, line 6 - line 13 column 15, line 19 - line 40	1-20
A	EP 0 573 398 A (HUGHES AIRCRAFT CO) 8 December 1993 (1993-12-08) page 2, line 53 -page 3, line 3 page 5, line 50 -page 6, line 12 figures 8,10	1-20
A	EP 0 631 274 A (AT & T CORP) 28 December 1994 (1994-12-28) abstract column 8, line 19 - line 32	1-20
A	PATENT ABSTRACTS OF JAPAN vol. 1999, no. 12, 29 October 1999 (1999-10-29) -& JP 11 184500 A (FUJITSU LTD), 9 July 1999 (1999-07-09) abstract	1,7,16
P,A	WO 00 25303 A (SALAMI REDWAN ;VOICEAGE CORP (CA); BESSETTE BRUNO (CA); LEFEBRE RO) 4 May 2000 (2000-05-04) page 33, line 10 - line 25	1,7,16

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/18232

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9727578	A	31-07-1997	US 6018706 A	25-01-2000
EP 0926660	A	30-06-1999	JP 11259098 A	24-09-1999
EP 0573398	A	08-12-1993	US 5495555 A	27-02-1996
			AT 174146 T	15-12-1998
			CA 2096991 A,C	02-12-1993
			DE 69322313 D	14-01-1999
			DE 69322313 T	01-07-1999
			FI 932465 A	02-12-1993
			JP 2010708 C	02-02-1996
			JP 6035500 A	10-02-1994
			JP 7036118 B	19-04-1995
			NO 931974 A	02-12-1993
			US 5734789 A	31-03-1998
			US 5651026 A	22-07-1997
			US 5596676 A	21-01-1997
EP 0631274	A	28-12-1994	CA 2124713 A,C	19-12-1994
			DE 69420200 D	30-09-1999
			DE 69420200 T	06-07-2000
			ES 2137325 T	16-12-1999
			JP 7168597 A	04-07-1995
			US 5719993 A	17-02-1998
JP 11184500	A	09-07-1999	NONE	
WO 0025303	A	04-05-2000	AU 6455599 A	15-05-2000
			AU 6456999 A	15-05-2000
			AU 6457099 A	15-05-2000
			AU 6457199 A	15-05-2000
			WO 0025305 A	04-05-2000
			WO 0025298 A	04-05-2000
			WO 0025304 A	04-05-2000